Expression, purification, crystallization and preliminary X-ray crystallographic analysis of tomato β -galactosidase 4

Masahiro Eda,^a Megumi Ishimaru^b and Toshiji Tada^{a*}

Received 24 September 2014

Accepted 20 December 2014

^aGraduate School of Science, Osaka Prefecture University, Sakai, Osaka 599-8531, Japan, and ^bFaculty of Biology-Oriented Science and Technology, Kinki University, Kinokawa, Wakayama 631-8505, Japan. *Correspondence e-mail: tada@b.s.osakafu-u.ac.jp

Keywords: β -galactosidase; *Solanum lycopersicum*; plant cell wall; TBG4.

Plant β -galactosidases play important roles in carbohydrate-reserve mobilization, cell-wall expansion and degradation, and turnover of signalling molecules during ripening. Tomato β -galactosidase 4 (TBG4) not only has β -galactosidase activity but also has exo- β -(1,4)-galactanase activity, and prefers β -(1,4)-galactans longer than pentamers as its substrates; most other β -galactosidases only have the former activity. Recombinant TBG4 protein expressed in the yeast *Pichia pastoris* was crystallized by the sitting-drop vapour-diffusion method using PEG 10 000 as a precipitant. The crystals belonged to the orthorhombic space group $P2_12_12_1$, with unit-parameters $a = 92.82$, $b = 96.30$, $c = 159.26$ Å, and diffracted to 1.65 Å resolution. Calculation of the Matthews coefficient suggested the presence of two monomers per asymmetric unit ($V_M = 2.2$ Å³ Da⁻¹), with a solvent content of 45%.

1. Introduction

β -Galactosidases (EC 3.2.1.23) catalyze the release of terminal nonreducing β -D-galactosyl residues from carbohydrates, glycolipids and glycoproteins. These enzymes are classified into four glycoside hydrolase (GH) families: GH1, GH2, GH35 and GH42 (Carbohydrate-Active Enzymes database; <http://www.cazy.org>). β -Galactosidases in the GH1, GH2 and GH42 families have been found predominantly in microorganisms (Cubellis *et al.*, 1990; Kalnins *et al.*, 1983; Ohtsu *et al.*, 1998), whereas the enzymes in GH35 have been found in microorganisms (Zinin *et al.*, 2002), animals (Ohto *et al.*, 2012) and plants (Smith & Gross, 2000). A number of plant β -galactosidase genes belonging to GH35 have been found (Gantulga *et al.*, 2008; Smith & Gross, 2000; Tanthanuch *et al.*, 2008) and further grouped into at least four subfamilies based on amino-acid sequence and function (Ishimaru *et al.*, 2009). Plant β -galactosidases have been reported to play important roles in the metabolism of galactosyl conjugates during carbohydrate-reserve mobilization, cell-wall expansion and degradation, and turnover of signalling molecules during ripening (de Alcântara *et al.*, 2006; Esteban *et al.*, 2003).

Tomato fruit development is marked by significant changes in the β -galactosyl content of the pericarp cell wall (Gross & Wallner, 1979; Kim *et al.*, 1991). Physiological studies suggest that a fourfold to fivefold increase in exo- β -galactanase activity coupled with a corresponding increase in free galactose may be involved in β -galactosyl loss during ripening (Carey *et al.*, 1995; Kim *et al.*, 1991; Pressey, 1983; Smith *et al.*, 2002). At least seven tomato β -galactosidases (TBG1–TBG7) belonging to GH35 have been suggested to play important roles during fruit development and maturation (Smith &

Table 1
Macromolecule-production information.

Source organism	<i>S. lycopersicum</i>
DNA source	cDNA
Forward primer	TGAATTCAGTGTTCCTTATGATGACAGAGC
Reverse primer	TCCCCGCGGATCTTGATCTCCTGACTAGAG
Cloning vector	pPICZα A
Expression vector	pPICZα A
Expression host	<i>P. pastoris</i> SMD1168H
Complete amino-acid sequence of the construct produced	EAEAEFSVSYYDDRAIIINGKRKILISGSIHYP-RSTPQMWPDLIQAKDGGDLVIETVVFVWNG-HEPSPGKYNFEGRYDLVRFIKMVQRAGLYV-NLRIGPYVCAEWNFGGFPVWLKYVPGMEFR-TNNQPFKVMQGFVQKIVNMMKSENLFESQ-GGPIIMAQIENEYGPVEWIGAPGKAYTKW-AAQMAVGLKTGVPWIMCKQEDAPDPVIDTC-NGFYCEGFRPNKPYKPKMWTEVWTGWYTKF-GGPIQRPAAEDIAFSVARFVQNNGSFFNY-MYHGNTNFGRTSSGLFIATSYDYDAPLDEY-GLLNEPKYGHRLDLHKAIKLSEPALVSSYA-AVTSLSGNSQEAHVYRSKSGACAFLSNYDS-RYSVKVTFQNRPNLPPWSISILPCKTAV-YNTAQVNSQSSSIKMTAGGGLSWQSYNEE-TPTADDSDTLTANGLWEQKNVTRDSSDYLW-YMTNVNIIASNEGFLKNGKDPYLTVMASGHV-LHVFVNGKLSGTIVYGTLDNPKLTYSGNVKL-RAGINKISLLSVSVGLPNVGVHYDTWNAAGV-LGPVTLTSLNEGSRNLAKQKWSYKVLKGE-SLSLHSLSGSSSVWVRGSLMAQKPLTWY-KATFNAPGGNDPLALDMSMGKGGIWIINGE-GVGRHWPQYIAQGDCKSKSYAGTFNEKKCCQ-TNCGQPSQRWYHVPRSWLKPSGNLLVVFEE-WGNGPTGISLVRRSRSAASAFLEQKLISE-EDLNSAVDHHHHHH

Gross, 2000). The antisense down-regulation of the TBG4 gene indicated that TBG4 is closely related to fruit softening (Smith *et al.*, 2002). TBG4 (724 amino acids) is the only enzyme belonging to the β -galactosidase/exo- β -(1,4)-galactanase subfamily among the tomato β -galactosidases (TBGs) and shows a strong ability to attack β -(1,4)-galactan; it works most efficiently when it can bind a substrate of at least five galactose residues (Ishimaru *et al.*, 2009).

Crystal structures have been elucidated for GH35 β -galactosidases isolated from various sources, including *Penicillium* sp. (1011 amino acids; PDB entry 1tg7; Rojas *et al.*, 2004), *Hypocrea jecorina* (1023 amino acids; PDB entry 3og2; Maksimainen *et al.*, 2011), *Homo sapiens* (677 amino acids; PDB entry 3thc; Ohto *et al.*, 2012), *Streptococcus pneumoniae* (595 amino acids; PDB entry 4e8c; Cheng *et al.*, 2012), *Aspergillus oryzae* (1005 amino acids; PDB entry 4iug; Maksimainen *et al.*, 2013), *Bacteroides thetaiotaomicron* (625 amino acids; PDB entry 3d3a; New York SGX Research Center for Structural Genomics, unpublished work) and *Caulobacter crescentus* (549 amino acids; PDB entry 3u7v; Midwest Center for Structural Genomics, unpublished work). These enzymes have the catalytic TIM-barrel domain in common. Although the catalytic domain of TBG4 displays moderate (27–34%) amino-acid sequence identity to these enzymes, the other part of TBG4 displays low (19–25%) amino-acid sequence identity to these enzymes. No β -galactosidase structures from plants have so far been reported. The clarification of the three-dimensional structure of TBG4 would allow elucidation of the catalytic function and the

substrate specificity. Here, we report the key steps, including expression, purification, crystallization and preliminary X-ray analysis, towards the structural characterization of TBG4.

2. Materials and methods

2.1. Protein production

The cDNA fragment encoding TBG4 (residues 24–724) was inserted downstream of the *Saccharomyces cerevisiae* α -factor signal sequence in the *Pichia pastoris* expression vector pPICZα A (Invitrogen) to produce a secreted recombinant protein fused to a hexahistidine tag at the C-terminus. Subsequently, this vector was transformed into *P. pastoris* SMD1168H strain by electroporation (Table 1) and Zeocin-resistant transformants were selected on YPDS+Zeocin agar plates (1% yeast extract, 2% peptone, 2% dextrose, 1 M D-sorbitol, 100 µg ml⁻¹ Zeocin, 2% agar). The transformants were grown on BMMY+X-gal agar plates (1% yeast extract, 2% peptone, 2% casamino acids, 1.34% yeast nitrogen base without amino acids, 0.1 M potassium phosphate buffer pH 6.0, 0.2 M D-sorbitol, 4 × 10⁻⁵% biotin, 50 µg ml⁻¹ 5-bromo-4-chloro-3-indolyl- β -D-galactoside, 0.5% methanol, 2% agar) for several days at 293 K, during which time 100 µl of methanol was added to the lids of each inverted plate every 24 h to induce expression of the recombinant protein. A blue colony was inoculated into YPG medium (1% yeast extract, 2% peptone, 1% glycerol) and grown at 303 K with shaking for 1 d. The cells were harvested by centrifugation and resuspended at an OD_{600 nm} of 1.0 in BMMY medium (1% yeast extract, 2% peptone, 2% casamino acids, 1.34% yeast nitrogen base without amino acids, 0.1 M potassium phosphate buffer pH 7.0, 0.2 M D-sorbitol, 4 × 10⁻⁵% biotin, 1.0% methanol) to induce expression. The yeast was cultured for

Figure 1
SDS-PAGE analysis of TBG4 purification. Lanes M, molecular-mass standards (labelled in kDa); lane A, TBG4 after nickel-affinity purification; lane B, TBG4 after enzymatic treatment and size-exclusion chromatography. Proteins were detected by staining with Coomassie Brilliant Blue.

Table 2
Crystallization.

Method	Sitting-drop vapour diffusion
Plate type	96-well protein crystallization plate (VIOLAMO)
Temperature (K)	277
Protein concentration (mg ml ⁻¹)	8
Buffer composition of protein solution	10 mM sodium acetate buffer pH 4.5 containing 200 mM L-arginine hydrochloride
Composition of reservoir solution	0.1 M HEPES buffer pH 7.5, 20% (w/v) PEG 10 000
Volume and ratio of drop	2.0 µl (1:1)
Volume of reservoir (µl)	50

24 h at 303 K and for 48 h at 293 K in a shaking incubator. Methanol was added to a final concentration of 1% every 24 h during the period of induction.

The medium containing TBG4 was centrifuged to remove the cells. Ammonium sulfate was added to the resulting supernatant to 75% saturation. The precipitate was collected by centrifugation and dissolved in 10 ml binding buffer (20 mM sodium phosphate buffer pH 8.0, 500 mM NaCl, 40 mM imidazole). The resulting solution was loaded onto a nickel-affinity column (Ni Sepharose 6 Fast Flow, GE Healthcare) equilibrated with the binding buffer and eluted with 50 mM sodium acetate buffer pH 4.5 containing 500 mM NaCl. The eluate was dialysed against 10 mM sodium acetate buffer pH 4.5 containing 0.2 M L-arginine hydrochloride and was incubated with α -mannosidase from *Canavalia ensiformis* (Sigma–Aldrich) for 3 d at 293 K to cleave polysaccharides from TBG4. Subsequently, the pH of the deglycosylated TBG4 solution was adjusted to 7.0 with 1 M Tris and the protein solution was then incubated with carboxypeptidase A (Sigma–Aldrich) for 1 d at 293 K to remove the C-terminal histidine tag and part of the linker peptide. The resulting enzyme solution was concentrated using Amicon Ultra Centrifugal Filters (Millipore) and applied onto a size-exclusion chromatography column (Superdex 200 10/300 GL, GE Healthcare) equilibrated with 10 mM sodium acetate buffer pH 4.5 containing 0.2 M L-arginine hydrochloride and eluted with the same buffer. Finally, the fraction showing a

Figure 2
The crystal of TBG4 in a cryoloop used for data collection. The dimensions of the crystal are $0.6 \times 0.3 \times 0.1$ mm.

single band on SDS–PAGE was concentrated to 8 mg ml^{-1} using Amicon Ultra Centrifugal Filters (Millipore) prior to crystallization. Protein purity was confirmed by SDS–PAGE analysis (Fig. 1).

2.2. Crystallization

Crystallization trials were performed using the sitting-drop vapour-diffusion method at 277 K. Crystallization conditions were initially screened using the commercially available screening kit Crystal Screen HT (Hampton Research). A 2.0 µl drop of TBG4 solution was mixed with an equal volume of crystallization solution and equilibrated against 1 ml crystallization solution. Using a condition consisting of 0.1 M HEPES buffer pH 7.5, 20% (w/v) PEG 10 000 (Table 2), crystals grew to maximal dimensions of approximately $0.6 \times 0.3 \times 0.1$ mm in three months (Fig. 2).

2.3. Data collection and processing

Prior to data collection, a crystal was soaked in cryoprotectant solution [crystallization solution supplemented with 25% (v/v) PEG 400]. The crystal was mounted in LithoLoops (Molecular Dimensions) and flash-cooled directly in a nitrogen-gas stream at 100 K. X-ray diffraction images were collected on a Quantum 315r CCD detector (ADSC) at the PF-17A station of the Photon Factory (PF), Ibaraki, Japan using synchrotron radiation of wavelength $0.980 \text{ \AA}$, an oscillation angle of 0.5° , an exposure time of 1 s per image, a rotation range of 180° and a crystal-to-detector distance of 220 mm (Fig. 3). The intensity data were processed and scaled using the *HKL-2000* package (Otwinowski & Minor, 1997).

Figure 3
X-ray diffraction pattern from the TBG4 crystal (oscillation width of 0.5°). Circles are labelled with the resolution limits in $\text{\AA}$. The inset is a magnification showing diffraction spots at $1.9 \text{ \AA}$ resolution.

Table 3

Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	BL17A, PF
Wavelength (Å)	0.980
Temperature (K)	100
Detector	Quantum 315r CCD
Crystal-to-detector distance (mm)	220
Rotation range per image (°)	0.5
Total rotation range (°)	180
Exposure time per image (s)	1
Space group	$P2_12_12_1$
a, b, c (Å)	92.82, 96.30, 159.26
Resolution range (Å)	50.0–1.65 (1.68–1.65)
Total No. of reflections	1271036
No. of unique reflections	170809
Completeness (%)	99.9 (99.5)
Multiplicity	7.4 (7.4)
$\langle I/\sigma(I) \rangle$	33.0 (3.8)
$R_{\text{meas}}^\dagger$	0.079 (0.53)
$R_{\text{p.i.m.}}^\ddagger$	0.029 (0.19)
Overall B factor from Wilson plot (Å ²)	23.3

$^\dagger R_{\text{meas}} = \frac{\sum_{hkl} \{N(hkl)/[N(hkl) - 1]\}^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$.
 $^\ddagger R_{\text{p.i.m.}} = \frac{\sum_{hkl} \{1/[N(hkl) - 1]\}^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$ (Weiss & Hilgenfeld, 1997).

3. Results and discussion

TBG4 with a C-terminal hexahistidine tag (736 amino-acid residues, 82 kDa) was expressed in *P. pastoris* and secreted into the culture medium. The recombinant protein in the culture supernatant was concentrated using an ammonium sulfate precipitation method and was purified using a nickel-affinity column. The purified protein appeared as a broad smeared band on SDS–PAGE (Fig. 1), which suggested that the recombinant TBG4 was highly glycosylated. Polysaccharide moieties bound to TBG4 were trimmed off by incubation with α -mannosidase and the C-terminal tag of TBG4 was then removed by incubation with carboxypeptidase A. TBG4 treated with these enzymes was purified by size-exclusion chromatography. SDS–PAGE showed that the solution mostly contained deglycosylated TBG4 (Fig. 1). The initial screening produced protein crystals in only one condition: 0.1 M HEPES buffer pH 7.0, 20% (w/v) PEG 10 000. The crystal grew to maximum dimensions of $0.6 \times 0.3 \times 0.1$ mm in the initial condition (Fig. 2). Diffraction data were collected from a crystal to 1.65 Å resolution on beamline BL17A at the Photon Factory using synchrotron radiation (Table 3). The crystal belonged to space group $P2_12_12_1$, with unit-cell parameters $a = 92.82$, $b = 96.30$, $c = 159.26$ Å. Calculation of the Matthews coefficient (Matthews, 1968) suggested the presence of two monomers per asymmetric unit ($V_M = 2.2$ Å³ Da^{−1}) with a solvent content of 45%. Recently, we have succeeded in determining the structure of TBG4 by the native sulfur-

based single-wavelength anomalous diffraction (native SAD) method with a molecular-replacement partial structure using a diffraction data set collected at a wavelength of 2.0 Å.

Acknowledgements

The synchrotron-radiation experiments were performed at the Photon Factory under Proposal No. 2011G144 and at SPring-8 under Proposal No. 2011B1499. We thank the beamline staff at BL-17A of PF, Tsukuba, Japan and at BL38B1 of SPring-8, Harima, Japan for providing data-collection facilities and support.

References

- Alcántara, P. H. N. de, Martim, L., Silva, C. O., Dietrich, S. M. & Buckeridge, M. S. (2006). *Plant Physiol. Biochem.* **44**, 619–627.
- Carey, A. T., Holt, K., Picard, S., Wilde, R., Tucker, G. A., Bird, C. R., Schuch, W. & Seymour, G. B. (1995). *Plant Physiol.* **108**, 1099–1107.
- Cheng, W., Wang, L., Jiang, Y.-L., Bai, X.-H., Chu, J., Li, Q., Yu, G., Liang, Q.-L., Zhou, C.-Z. & Chen, Y. (2012). *J. Biol. Chem.* **287**, 22910–22918.
- Cubellis, M. V., Rozzo, C., Montecucchi, P. & Rossi, M. (1990). *Gene*, **94**, 89–94.
- Esteban, R., Dopico, B., Muñoz, F. J., Romo, S., Martín, I. & Labrador, E. (2003). *Plant Cell Physiol.* **44**, 718–725.
- Gantulga, D., Turan, Y., Bevan, D. R. & Esen, A. (2008). *Phytochemistry*, **69**, 1661–1670.
- Gross, K. C. & Wallner, S. J. (1979). *Plant Physiol.* **63**, 117–120.
- Ishimaru, M., Smith, D. L., Mort, A. J. & Gross, K. C. (2009). *Planta*, **229**, 447–456.
- Kalnins, A., Otto, K., Rüther, U. & Müller-Hill, B. (1983). *EMBO J.* **2**, 593–597.
- Kim, J., Gross, K. C. & Solomos, T. (1991). *Postharvest Biol. Technol.* **1**, 67–80.
- Maksimainen, M. M., Hakulinen, N., Kallio, J. M., Timoharju, T., Turunen, O. & Rouvinen, J. (2011). *J. Struct. Biol.* **174**, 156–163.
- Maksimainen, M. M., Lampio, A., Mertanen, M., Turunen, O. & Rouvinen, J. (2013). *Int. J. Biol. Macromol.* **60**, 109–115.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Ohto, U., Usui, K., Ochi, T., Yuki, K., Satow, Y. & Shimizu, T. (2012). *J. Biol. Chem.* **287**, 1801–1812.
- Ohtsu, N., Motoshima, H., Goto, K., Tsukasaki, F. & Matsuzawa, H. (1998). *Biosci. Biotechnol. Biochem.* **62**, 1539–1545.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Pressey, R. (1983). *Plant Physiol.* **71**, 132–135.
- Rojas, A. L., Nagem, R. A., Neustroev, K. N., Arand, M., Adamska, M., Eneyskaya, E. V., Kulminkaya, A. A., Garratt, R. C., Golubev, A. M. & Polikarpov, I. (2004). *J. Mol. Biol.* **343**, 1281–1292.
- Smith, D. L., Abbott, J. A. & Gross, K. C. (2002). *Plant Physiol.* **129**, 1755–1762.
- Smith, D. L. & Gross, K. C. (2000). *Plant Physiol.* **123**, 1173–1183.
- Tanthanuch, W., Chantarangsee, M., Maneesan, J. & Ketudat Cairns, J. R. (2008). *BMC Plant Biol.* **8**, 84.
- Weiss, M. S. & Hilgenfeld, R. (1997). *J. Appl. Cryst.* **30**, 203–205.
- Zinin, A. I., Eneyskaya, E. V., Shabalin, K. A., Kulminkaya, A. A., Shishlyannikov, S. M. & Neustroev, K. N. (2002). *Carbohydr. Res.* **337**, 635–642.